

# Energy Aware Architecture for Coordinated Mobility: An Approximate Dynamic Programming Approach

Hassan Jaleel and Jeff S. Shamma

## Abstract

Our goal is to design distributed coordination strategies that enable agents to achieve global performance guarantees while minimizing the energy cost of their actions with an emphasis on feasibility for real-time implementation. As a motivating scenario that illustrates the importance of introducing energy awareness at the agent level, we consider a team of mobile nodes that are assigned the task of establishing a communication link between two base stations with minimum energy consumption. We formulate this problem as a dynamic program in which the total cost of each agent is the sum of both mobility and communication costs. To ensure that the solution is distributed and real time implementable, we propose multiple suboptimal policies based on the concepts of approximate dynamic programming. To provide performance guarantees, we compute upper bounds on the performance gap between the proposed suboptimal policies and the global optimal policy. Finally, we discuss merits and demerits of the proposed policies and compare their performance using simulations.

## I. INTRODUCTION

The goal of this paper is to present a framework for multiagent systems that will allow individual agents to strike a balance between global performance and the cost of distributedly computing the globally optimal control action. This computation cost typically comprises the cost of communication among neighboring nodes. In the proposed framework, the focus is on designing energy aware local interaction laws for individual agents that can provide performance guarantees, are implementable in real time, and have limited communication and computation overhead. The framework is presented in the context of a motivating example in which a collection of relay nodes is deployed to establish a communication link between two base stations for a long period of time. The task is to find an energy efficient and distributed mobility strategy to move the relay nodes to optimal locations such that the total energy consumption is minimized. We will formulate this problem as a dynamic program in which the cost to

H. Jaleel and J.S. Shamma are with the King Abdullah University of Science and Technology (KAUST), Computer, Electrical and Mathematical Sciences and Engineering Division (CEMSE), Thuwal 23955–6900, Saudi Arabia. Email: hassan.jaleel@kaust.edu.sa, jeff.shamma@kaust.edu.sa. Shamma is also with the Georgia Institute of Technology, School of Electrical and Computer Engineering. Research supported by funding from KAUST.

be minimized is the sum of the communication and the mobility costs of all the agents. To ensure that the control policy is distributed, energy efficient, and real time implementable, we will propose multiple suboptimal policies and compare their performance with the global optimal policy based on the techniques of approximate dynamic programming [1].

A fundamental challenge in designing an energy aware scheme for multiagent systems is that each agent only has limited information of the system but its actions have direct impact on the global performance (see [2] and the references therein). Another challenge is to ensure that the proposed scheme is feasible for real-time implementation. This challenge gets more complicated when the objective is to find an optimal trajectory over a given interval because each agent must account for all the possible future trajectories of its neighbors while computing its current control actions, which makes real time implementation impractical even for simple scenarios. Consequently, the control strategies either become too complex and/or require excessive communication among the agents, resulting in large energy consumption. Therefore, an efficient energy aware coordination strategy must have the capability to strike a balance between the performance requirements of the system and the energy requirements of the control strategies that it implements to achieve this performance.

The importance of introducing energy awareness in multiagent systems like wireless sensor networks is widely recognized (see [3] and the references therein). However, when it comes to distributed energy aware mobility strategies, the existing literature is somewhat limited and there are still a lot of unanswered questions. In [4], a distributed energy aware coverage scheme was proposed but it only considered the tradeoff between sensing and processing and did not assign any cost to mobility. In [5], synchronous and asynchronous distributed algorithms were presented for steering relay nodes to the optimal locations for establishing a communication link between two base stations. However, no cost was assigned to mobility because of the assumption that either the batteries can be recharged or communication is for long time and mobility cost is negligible in comparison with communication cost. In [6], a similar problem was considered under a static setting in which agents initially determine and move to their optimal locations and then the communication starts. The proposed solutions were based on heuristics and no optimality guarantees were provided. In [7] and [8] both mobility and communication costs were considered and the problem was formulated as an optimal control problem in a dynamic setting in which agents move and communicate at the same time. In both the references, only a centralized setup was considered. References [9], [10], [11], [12], [13], and [14] investigated optimal consensus problems using model predictive control (MPC). In all of these works, different approximate solutions were proposed with guaranteed asymptotic convergence to the consensus set, but none of these works analyzed the energy consumption profile of the proposed solutions.

In this work, we will start by formulating the problem under consideration as an infinite horizon discounted LQR problem. Using the principal of optimality, we will formulate an equivalent one stage lookahead problem with optimal cost to go as terminal cost. This terminal cost will be a quadratic function of a positive definite matrix  $K$  such that  $K_{ij} \neq 0$  will indicate that the terminal cost of node  $i$  depends on node  $j$ . We will show that all the entries of  $K$  are non-zero which implies that the terminal cost of each node depends on the states of all the nodes. For a system with fixed communication network, the matrix  $K$  can be computed offline before the system is deployed.

Therefore, the original dynamic program will be reduced to a parameter optimization problem over a network of agents for which distributed optimization algorithms exist that can guarantee global optimal solution (see e.g., [15] and [16]).

In [15], a dual decomposition based distributed optimization algorithm was presented in which dual variables were introduced to decouple the cost of each agent. However, each agent needs to communicate with all the agents whose state directly influence its cost. Since all the entries of  $K$  are non-zeros, the communication network will be a complete graph and the problem will become centralized. In [16], a consensus based algorithm was presented in which only neighboring nodes of a graph were required to communicate as long as the graph was connected. However, the size of the message that agents have to communicate with their neighbors directly depends on the number of agents which can be large and can result in significant energy consumption.

*In the proposed framework, we will impose the constraints of the communication network on the optimal cost to go function and will propose approximate cost to go functions such that the cost of each agent will only depend on itself and its neighboring nodes.* This will significantly reduce the size of the communication message since the number of neighboring nodes is typically small as compared to the total nodes in the network. The price of approximating the optimal cost to go will be a loss in the global performance. We will analyze this performance loss and provide upper bounds on the performance gap between the actual and optimal performances.

The remainder of this paper is organized as follows. Section II motivates the problem under investigation and presents a mathematical formulation. Section III presents the proposed schemes along with their performance bounds, which are the main results of this paper. Section IV provides performance comparisons of the proposed schemes based on simulations. Finally, Section V concludes the paper.

## II. PROBLEM DESCRIPTION

### A. Notation

We denote a graph by  $G = (V, E)$  where  $V = \{1, 2, \dots, N\}$  is the set of vertices and  $E = \{(i, j) \mid i, j \in V\}$  is the set of edges. Graph  $G$  is undirected if its links are bidirectional ( $(i, j) \in E$  iff  $(j, i) \in E$ ). The neighborhood set of node  $i$  is

$$\mathcal{N}_i = \{j \in V \mid (i, j) \in E\},$$

and the cardinality of this set is  $|\mathcal{N}_i|$ . A graph is connected if given any pair of nodes  $(i, j)$ , either  $(i, j) \in E$  or there exist some intermediate nodes such that  $\{(i, i_1), (i_1, i_2), \dots, (i_m, i_j)\} \in E$ . The degree matrix of  $G$  is a diagonal matrix  $\mathcal{D}(G)$  with the diagonal entries  $\mathcal{D}_{ii}(G) = |\mathcal{N}_i|$ . The adjacency matrix  $\mathcal{A}(G)$  is

$$\mathcal{A}_{ij}(G) = \begin{cases} 1 & : j \in \mathcal{N}_i \\ 0 & : \text{otherwise} \end{cases}$$

The graph laplacian,

$$\mathcal{L}(G) = \mathcal{D}(G) - \mathcal{A}(G),$$

is a symmetric and positive semidefinite matrix with real and non-negative eigenvalues for undirected graphs. Moreover, for a connected graph,  $0 = \lambda_1 < \lambda_2 \leq \lambda_3 \leq \dots \leq \lambda_N$ .

Let  $z_i(k) = [x_i(k) \ y_i(k)]^T$  denotes the location of node  $i$  at time  $k$  in  $\mathbb{R}^2$ , where  $[\cdot]^T$  is the transpose of a vector. For concise notation, the locations of all the nodes at time  $k$  are stacked in vector  $\mathbf{z}_k \in \mathbb{R}^{2N}$ , i.e.,  $\mathbf{z}_k = [z_1(k)^T \ \dots \ z_N(k)^T]^T$ . For  $i \in \{1, 2, \dots, N\}$ ,  $\mathbf{z}_{-i,k}$  denotes the vector of locations of all neighbors of  $i$ , i.e.,  $\mathbf{z}_{-i,k} = [z_{i_1}(k)^T \ z_{i_2}(k)^T \ \dots \ z_{i_{|\mathcal{N}_i|}}(k)^T]^T$  where  $i_1, \dots, i_{|\mathcal{N}_i|} \in \mathcal{N}_i$ . For  $z \in \mathbb{R}^2$ ,  $\|z\|$  denotes the Euclidean norm.

For a matrix  $A$  of dimensions  $N \times M$ ,  $\|A\|_F$  is the Frobenius norm, i.e.,

$$\|A\|_F = \sqrt{\sum_{i=1}^N \sum_{j=1}^M |A_{ij}|^2}.$$

If  $A \in \mathbb{R}^{N \times N}$  and  $\lambda_1 \geq \lambda_2 \geq \dots \lambda_N$  are its  $N$  eigenvalues in non-increasing order, then  $\lambda_i(A) = \lambda_i$  which implies that  $\lambda_i$  will be used both as a function and as a constant value. The vector  $\mathbf{1}_n$  denotes the vector  $[1 \ 1 \ \dots \ 1]^T \in \mathbb{R}^n$ ,  $I_n \in \mathbb{R}^{n \times n}$  denotes identity matrix, and  $0_{n \times m}$  denotes a matrix of dimension  $n \times m$  with all entries equal to 0. Boldface letters like  $\mathbf{z}_k$  denote collections of vectors, and their subscript represents time.

### B. Definitions

Let  $A$  be an  $N \times N$  matrix. Then the spectrum of  $A$  is the set of all the eigenvalues of  $A$  and is denoted as  $\text{spec}(A)$ . The spectral radius of  $A$  is  $\rho(A) = \max\{|\lambda_1|, |\lambda_2|, \dots, |\lambda_N|\}$ .

*Definition 2.1:* A matrix  $D$  is a positive matrix ( $D > 0$ ) if all of its entries are positive. It is a non-negative matrix ( $D \geq 0$ ) if all of its entries are non-negative.

*Definition 2.2:* An  $N \times N$  matrix  $Z$  is an  $M$  matrix if it can be written as

$$Z = sI_N - D, \quad s > 0, \ D \geq 0, \ \rho(D) \leq s,$$

If  $\rho(D) < s$  then  $Z$  is a non-singular  $M$  matrix and if  $\rho(D) = s$  then  $Z$  is a singular  $M$  matrix. An essential property of an  $M$  matrix is that  $Z_{ii} > 0$  and  $Z_{ij} \leq 0$  for all  $i$  and  $j$ . Moreover, the inverse of an  $M$  matrix is a strictly positive matrix, i.e., all entries are positive [17].

*Gershgorin circle theorem:* If  $A$  is an  $N \times N$  matrix, then every eigenvalue of  $A$  lies in at least one of the circles  $C_1, C_2, \dots, C_N$  where  $C_i$  has its center at the diagonal entry  $c_i(A) = A_{ii}$  and its radius  $r_i(A) = \sum_{j \neq i} |A_{ij}|$ . The Gershgorin circle theorem provides upper and lower bounds for the spectrum of a square matrix. For a laplacian matrix  $\mathcal{L}$ ,  $c_i(\mathcal{L}) = r_i(\mathcal{L})$  since each diagonal entry is equal to the sum of the absolute values of the off diagonal entries of that row.

### C. System Setup

Consider two base stations  $\mathcal{B}_1$  and  $\mathcal{B}_2$  separated by a distance  $d$ . Without loss of generality we can assume that the base station  $\mathcal{B}_1$  is located at the origin. The objective is to establish an uninterrupted communication link between these base stations for a time interval of length  $T$  with minimum energy consumption. From [5], the power

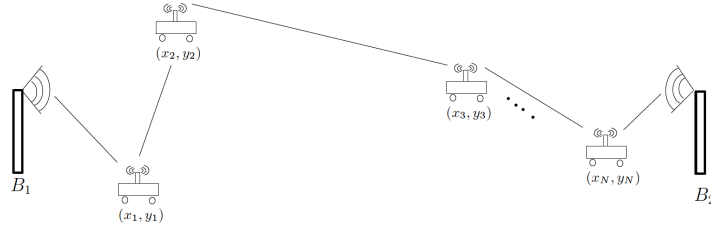


Fig. 1. Illustration of the communication network.  $B_1$  and  $B_2$  are the base stations and  $N$  relay nodes are located at  $(x_i, y_i)$  for all  $i \in \{1, 2, \dots, N\}$ .

required to successfully transmit over a distance  $d$  at maximum data rate is directly proportional to the square of the transmission distance  $d$ , i.e.,

$$P_{\text{comm}} = (\kappa_c d^2 + b),$$

where  $\kappa_c$  is a proportionality constant and  $b$  reflects the additional power consumed by transmitter/receiver circuitry. In this work we are interested in communication power only, so assume  $b = 0$ .

The energy required to establish the communication link between the two base stations will be extremely high for large values of  $d$  and  $T$ . One solution is to use  $N$  relay nodes as shown in Fig. 1. Depending on the value of  $N$  and the deployment locations of the nodes, the overall energy consumption can be significantly reduced. For minimum energy consumption, the relay nodes must be evenly spaced on the straight line between the two base stations [5]. However, we are interested in a scenario in which the nodes are randomly deployed between the base stations and they need to estimate and move to the optimal deployment locations in a decentralized manner while communicating. This redeployment towards the optimal locations will have its own cost that must be accounted for.

Let  $z_i(k)$  denotes the location of relay node  $i$  at time  $k$  for  $i \in \{1, 2, \dots, N\}$ , and  $z_{N+1}(k)$  and  $z_{N+2}(k)$  denote the locations of the base stations  $B_1$  and  $B_2$  respectively. Because the base stations remain stationary,  $z_{N+1}(k)$  and  $z_{N+2}(k)$  are constant values. This network of the relay nodes and the base stations is represented by a graph  $G_c = (V, E)$  in which the vertex set  $V$  consists of  $N$  relay nodes  $\{1, 2, \dots, N\}$  and two base stations  $\{N+1, N+2\}$ . An edge exists between two vertices in  $G_c$  if the corresponding nodes can communicate with each other, i.e.,  $\mathcal{N}_i$  is the set of nodes with which node  $i$  can communicate. This communication network can be represented algebraically by graph laplacian  $\mathcal{L} = \mathcal{L}(G_c)$  where

$$\mathcal{L}_{ij} = \begin{cases} -1 & : i \neq j \text{ and } j \in \mathcal{N}_i \\ |\mathcal{N}_i| & : i = j \\ 0 & : \text{otherwise} \end{cases}$$

To find the optimal locations of the relay nodes, we assume that the system satisfies the following properties.

*Assumptions:*

- 1) The communication network is a line graph, and it remains fixed, i.e.,  $\mathcal{L} \in \mathbb{R}^{(N+2) \times (N+2)}$  and has the following structure:

$$\mathcal{L} = \begin{pmatrix} 2 & -1 & 0 & \cdots & \cdots & -1 & 0 \\ -1 & 2 & -1 & \cdots & \cdots & 0 & 0 \\ 0 & -1 & 2 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & -1 & 2 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 2 & 0 & -1 \\ -1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 1 \end{pmatrix}$$

- 2) The relay nodes are mobile nodes with single integrator dynamics

$$z_i(k+1) = z_i(k) + u_i(k),$$

where  $u_i(k) \in \mathbb{R}^2$  is the input at time  $k$ .

Although the communication link is established initially, the overall energy consumption can be minimized if the relay nodes can somehow move to the optimal locations in an efficient manner. However, there is no leader or centralized authority that has the knowledge of the optimal locations, so the nodes need to figure it out locally based on communication with their neighbors. To ensure that the overall energy consumption is minimized, it is imperative to incorporate the cost of mobility in the system model because mobility is orders of magnitude more expensive than communication. From [18], mobility power consumption can be approximated to be a function of speed, i.e.,

$$P_{\text{mob}} = \psi(\|u(k)\|),$$

where  $u(k)$  is the robot's velocity at time  $k$ . We assume that mobility cost of a node is proportional to the square of its speed, i.e.,  $\psi(\|u(k)\|) = \kappa_M \|u(k)\|^2$ , where  $\kappa_M$  is a mobility constant. This choice of mobility model is valid for mobile nodes that use DC motors and operate at low speeds. The advantage of this model is that it will help in obtaining an analytical solution to the optimization problem.

In this paper, we will formulate this problem as a dynamic program and propose suboptimal policies that can be implemented in real time and have low energy overhead. We will also derive upper bounds on the performance gap between the proposed policies and the global optimal policy and compare their performance through simulations.

#### D. Problem Formulation

The total cost that is to be minimized is the sum of the mobility cost and the communication cost of the system. We will formulate the problem as an infinite horizon problem with a discount factor  $\alpha \in (0, 1)$ .

*Problem:*

$$\begin{aligned} \lim_{T \rightarrow \infty} \min_{\{\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_{T-1}\}} \sum_{k=0}^{T-1} \alpha^k g(\mathbf{z}_k, \mathbf{u}_k) + \alpha^T J_T(\mathbf{z}_T) \\ \text{s.t. } \mathbf{z}_{k+1} = \mathbf{z}_k + B\mathbf{u}_k, \end{aligned} \quad (\mathcal{P})$$

where

$$\begin{aligned} g(\mathbf{z}_k, \mathbf{u}_k) &= \mathbf{J}_{\text{comm}}(\mathbf{z}(k)) + \mathbf{J}_{\text{mob}}(\mathbf{u}_k) \\ \mathbf{J}_{\text{comm}}(\mathbf{z}_k) &= \frac{1}{2} \left( \sum_{i=1}^N \sum_{j \in \mathcal{N}_i} \kappa_C \|z_i(k) - z_j(k)\|^2 \right), \\ \mathbf{J}_{\text{mob}}(\mathbf{u}_k) &= \sum_{i=1}^N \kappa_M \|u_i(k)\|^2, \\ J_T(\mathbf{z}_T) &= \frac{1}{2} \left( \sum_{i=1}^N \sum_{j \in \mathcal{N}_i} \kappa_C \|z_i(T) - z_j(T)\|^2 \right). \end{aligned}$$

In the system dynamics,

$$B = \begin{bmatrix} I_{2N} \\ 0_{4 \times 2N} \end{bmatrix},$$

where  $0_{4 \times 2N}$  corresponds to the base stations that are stationary. The cost can be represented compactly in matrix vector notation as follows:

$$\begin{aligned} g(\mathbf{z}_k, \mathbf{u}_k) &= \mathbf{z}_k^T Q \mathbf{z}_k + \mathbf{u}_k^T R \mathbf{u}_k, \\ J_T(\mathbf{z}_T) &= \mathbf{z}_T^T Q_T \mathbf{z}_T. \end{aligned} \quad (1)$$

Here,  $R = \kappa_M I_{2N}$  is a diagonal matrix, and  $Q = \kappa_C (\mathcal{L} \otimes I_2)$  is a symmetric and positive semidefinite matrix of dimensions  $2(N+2) \times 2(N+2)$ . With linear dynamics and quadratic cost,  $\mathcal{P}$  is an LQR problem where  $u_i(k) = \mu_{i,k}(\mathbf{z}_k)$  is the control value of agent  $i$  located at  $z_i(k)$  at time  $k$ ,  $\mu_{i,k} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  is the policy of agent  $i$  at time  $k$ , and  $\boldsymbol{\mu}_k = [\mu_{1,k}^T, \mu_{2,k}^T, \dots, \mu_{N,k}^T]^T$  is a policy vector at time  $k$ .

Problem  $\mathcal{P}$  is a standard infinite horizon dynamic program with  $\boldsymbol{\mu}_k$  specific stage cost  $g(\mathbf{z}_k, \boldsymbol{\mu}_k(\mathbf{z}_k))$  and terminal cost  $J_T(\mathbf{z}_T)$ . Typically, dynamic programs suffer from the curse of dimensionality and become computationally intractable even for small size problems. However,  $\mathcal{P}$  is an LQR problem with closed form analytical solution that involves solving the following difference Riccati equation iteratively:

$$\begin{aligned} K_T &= Q_T \\ K_k &= (\alpha K_{k+1} - \alpha^2 K_{k+1} B (\alpha B^T K_{k+1} B + R)^{-1} B^T K_{k+1}) + Q, \end{aligned}$$

where  $K_k$  is symmetric and positive semidefinite for all  $k$ . The optimal cost to go at stage  $k$  is

$$J_k^*(\mathbf{z}) = \mathbf{z}^T K_k \mathbf{z}.$$

The advantage of formulating  $\mathcal{P}$  as an infinite horizon problem is that  $K_k \rightarrow K$  as  $T \rightarrow \infty$ :

$$K = \alpha K - \alpha^2 K B (\alpha B^T K B + R)^{-1} B^T K + Q. \quad (2)$$

where  $K$  is a symmetric and positive definite matrix. Because of the stationarity of  $K$ , the optimal cost is also stationary, i.e.,  $J_k^*(\mathbf{z}) \rightarrow J^*(\mathbf{z})$  as  $k \rightarrow \infty$ , where

$$J^*(\mathbf{z}) = \mathbf{z}^T K \mathbf{z}. \quad (3)$$

Using the principal of optimality,  $\mathcal{P}$  can be formulated as a one stage lookahead problem with stationary optimal cost to go as its terminal cost, i.e.,

$$\boldsymbol{\mu}^*(\mathbf{z}_k) = \arg \min_{\mathbf{u}} [g(\mathbf{z}_k, \mathbf{u}) + \alpha J^*(\mathbf{z}_k + B\mathbf{u})] \quad (4)$$

subject to the dynamics and constraints in  $\mathcal{P}$ . The optimal stationary policy  $\boldsymbol{\mu}^*$  has the form

$$\begin{aligned} \boldsymbol{\mu}^*(\mathbf{z}_k) &= -L\mathbf{z}_k, \\ L &= \alpha(\alpha B^T K B + R)^{-1} B^T K. \end{aligned} \quad (5)$$

The optimal policy is a simple static feedback law that can be implemented easily under normal circumstances. However, we will show that the optimal policy is centralized which is an intuitive result and makes this problem challenging since the optimal solution requires complete knowledge of the network and cannot be decentralized.

Next we will show that  $K$  is a non-singular  $M$ -matrix and all of its entries are non-zero, which means each agent requires state information of all the other agents to compute the optimal cost to go.

*Proposition 1:* If  $Q \in \mathbb{R}^{2(N+2) \times 2(N+2)}$  is a symmetric  $M$  matrix,  $B = \begin{bmatrix} I_{2N} \\ 0_{4 \times 2N} \end{bmatrix}$ , and  $R = \kappa_M I_{2N}$ , then the positive definite solution of the Riccati equation

$$K = \alpha K - \alpha^2 K B (\alpha B^T K B + R)^{-1} B^T K + Q$$

is also an  $M$  matrix. Moreover, if  $Q$  is a laplacian matrix of a connected graph then all the entries of  $K$  are non-zero.

*Proof:* The proof is presented in the appendix. ■

Problems having structure similar to  $\mathcal{P}$  have been studied in the context of optimal consensus problems for general LTI systems (see for example [9], [10], [12], [13], [14], and [19]). In all of these references, consensus problem was formulated as an infinite horizon LQR problem with the objective of minimizing disagreement among the agents with minimum control effort. In [19] it was shown that the optimal solution to this problem is centralized. Since the problem could not be solved in a decentralized manner, approximations were introduced to decouple the cost along the entire trajectory. In [9], [11], [13] and [14] each agent assumed that all of its neighbors remain stationary for all the future time. Based on this assumption, optimal control trajectories were computed and implemented for



one time step, a standard approach in MPC, and the process was repeated. In [10], the same stationarity assumption was used but the update rule was considered to be sequential, i.e., at each time only one agent was allowed to compute its optimal trajectory. In [12], each agent transmitted its assumed trajectory to all of its neighbors. At the same time it received such trajectories from its neighbors. Then, each agent used these assumed trajectories of its neighbors to solve its optimization problem and computed its actual trajectory. For stability, the actual trajectory and the assumed trajectory of each agent should be close.

Although  $\mathcal{P}$  has the same cost as the references mentioned, the goal is significantly different. The schemes in the references are primarily designed for achieving a certain steady state performance (consensus). To ensure that the trajectories computed by agents are feasible, and converge to the consensus set, these schemes rely on extensive communication among neighboring agents. In fact, some schemes require communication of entire trajectories among the neighboring agents repeatedly at each decision time step ([10] and [20]). This communication results in considerable energy overhead that is typically not taken into account.

*Our focus is on developing a framework for minimizing this communication overhead and ensuring that the scheme is distributed and real time implementable.* The price of reducing inter-agent communication will be in terms of global performance. Therefore, we will analyze the effects on system performance and provide upper bounds on the performance gap between the proposed and optimal policies. Furthermore, the proposed framework is not limited to a particular problem. Instead, it is intended to provide energy aware local coordination algorithms that can provide performance guarantees for a class of multiagent systems with fixed communication network.

#### E. Main Contribution

Based on the above discussion, it is inefficient in terms of communication energy overhead to solve problem  $\mathcal{P}$  in its original form. Another formulation of this problem is presented in Eq. (4) in which the problem is reduced to one step lookahead with optimal cost to go as the terminal cost. The optimal cost to go is a quadratic function of  $K$ , the positive definite solution of the Riccati equation (2), which is a function of communication network  $\mathcal{L}$ , system dynamics  $B$ , and communication and mobility constants  $\kappa_C$  and  $\kappa_M$  respectively. From Assumption (1),  $\mathcal{L}$  is fixed, so  $K$  can be computed offline before the nodes are deployed. From Prop. 1, all the entries of  $K$  are non-zero. Therefore, the cost in Eq. (4) has a local component and global component. Each node can compute its stage cost  $g(\mathbf{z}_k, \mathbf{u})$  using the current information of its neighbors only. However, the terminal cost requires information of all the nodes in the network.

Since both the stage and the terminal costs are convex, the distributed optimization algorithm presented in [16] can guarantee global optimal solution by allowing communication among neighboring nodes only. The algorithm is a consensus based distributed optimization algorithm to minimize the following cost

$$\min \sum_{i=1}^N f_i(x) \quad \text{s.t.} \quad x \in \mathbb{R}^m,$$

where  $x$  is the decision vector,  $f_i : \mathbb{R}^m \rightarrow \mathbb{R}$  is the local cost of agent  $i$ , and each agent can only communicate with its neighbors. To solve this problem, each agent maintains an estimate of the entire decision vector. It communicates

its estimate with its neighbors, receives their estimates, and updates its estimate of the decision vector by combining the information it received from its neighbors. For computing action at time  $k$ , the above process is repeated for a specified number of iterations which impacts the quality of the solution. If the size of the decision vector is small, the communication overhead of the algorithm can be tolerated. However, in multiagent systems, the size of the decision vector typically depends directly on the number of nodes  $N$  and can be very large. Consequently, the size of the data packet can get large which will result in significant energy overhead.

In the proposed framework, we approximate the global component of the cost (terminal cost) with a function that is close to the original function in some sense and is locally computable. This implies that the size of the data packet for each node will depend on the cardinality of its neighborhood set. Typically, multiagent systems are sparsely connected, i.e.,  $|\mathcal{N}_i| \ll N$ , so the size of the data packet is significantly reduced. *The main idea proposed in this work for designing suboptimal policies for each node  $i \in \{1, 2, \dots, N\}$  is to impose the communication network constraints of node  $i$  on the global optimal cost matrix  $K$ .* The resulting matrix  $K_{\mathcal{L}}$  will observe the constraints of the communication network, i.e.,  $(K_{\mathcal{L}})_{ij} = 0$  if  $j \notin \mathcal{N}_i$ , and will be used to design approximate cost to go for the LQR problem. The suboptimal policy for each node is to solve one step lookahead problem with this approximate cost to go as the terminal cost using some efficient distributed optimization algorithm. This procedure for designing suboptimal policies will guarantee that the terminal cost can be computed locally and will simplify performance comparison with the global optimal.

### III. ENERGY AWARE ARCHITECTURES

In this section we propose three Energy Aware Policies (EAP)s. The first two policies are distributed in nature and the last one is decentralized.

*Definition 3.1:* A policy  $\mu$  is distributed if the control action of each agent at time  $k$  is a function of its current state and the current state and input of its neighbors, i.e.,  $\mu_{i,k}$  depends on  $z_i(k)$ ,  $\mathbf{z}_{-i,k}$  and  $\mathbf{u}_{-i,k}$ .

*Definition 3.2:* A policy  $\mu$  is decentralized if the control action of each agent is a function of its current state and the current states of its neighbors, i.e.,  $\mu_{i,k}$  depends on  $z_i(k)$ ,  $\mathbf{z}_{-i,k}$ .

In a distributed policy, node  $i$  will have to repeatedly communicate with its neighbors to know their current inputs. However, the current state of the neighbors in a decentralized policy can either be sensed if the nodes are equipped with the required sensors, or it can be communicated by one time communication.

#### A. Distributed Energy Aware Policies

Next we propose two distributed energy aware policies EAPs I & II. In EAP I, we formulate a semidefinite program to find a matrix that is closest to  $K$  in terms of Frobenius norm and satisfies the desired sparsity structure.

### Energy Aware Policy I

Compute the positive definite matrix  $K$  offline that satisfies Riccati equation (2). Solve the following semidefinite program to find the projection of  $K$  on the sparsity structure of  $\mathcal{L}' = (\mathcal{L} \otimes I_2)$ .

$$\begin{aligned} & \min_{K_{\mathcal{L}}^1} \|K_{\mathcal{L}}^1 - K\|_F \\ & \text{s.t.} \quad K_{\mathcal{L}}^1 \geq 0, \\ & (K_{\mathcal{L}}^1)_{ij} = 0 \quad \text{if} \quad \mathcal{L}'_{ij} = 0 \\ & \forall i \text{ and } j \in \{1, \dots, 2(N+2)\}. \end{aligned}$$

The proposed approximate cost to go function is

$$\tilde{J}^1(\mathbf{z}_k) = \mathbf{z}_k^T K_{\mathcal{L}}^1 \mathbf{z}_k,$$

and each node solves the following one stage optimization problem to compute its stationary suboptimal policy:

$$\boldsymbol{\mu}^1(\mathbf{z}_k) = \arg \min_{\mathbf{u}} g(\mathbf{z}_k, \mathbf{u}) + \alpha \tilde{J}^1(\mathbf{z}_k + B\mathbf{u}), \quad (6)$$

with dynamics and constraints specified in  $\mathcal{P}$ .

*Proposition 2:* Let  $P_K$  be defined as

$$(P_K)_{ij} = \begin{cases} K_{ij} & : \mathcal{L}'_{ij} \neq 0 \\ 0 & : \text{otherwise} \end{cases}$$

Then  $K_{\mathcal{L}}^1 = P_K$ , i.e., the projection of  $K$  on  $\mathcal{L}'$  can be computed by simply replacing the entries of  $K$  corresponding to non-neighboring agents with zeros.

*Proof:* Firstly,  $P_K$  satisfies the desired sparsity constraints by construction. Secondly,  $P_K$  is symmetric because  $K$  and  $\mathcal{L}'$  are both symmetric. To show that  $P_K$  is positive definite, we use the fact that  $K$  is an  $M$  matrix, i.e., all the diagonal entries are positive and the off-diagonal entries are non-positive, and  $K_{ii} > -\sum_{j=1}^{N+2} K_{ij}$  since  $K > 0$ . Therefore, all the Gershgorin circles of  $P_K$  lie in the positive half plane, and so all of its eigenvalues are positive. Thus,  $P_K$  satisfies all the constraints. To prove that  $P_K$  is optimal, we use the definition of Frobenius norm.

$$\|P_K - K\|_F = \sqrt{\sum_{i=1}^{2(N+2)} \sum_{j=1}^{2(N+2)} |(P_K)_{ij} - K_{ij}|^2}$$

For  $(i, j)$  pair such that  $\mathcal{L}'_{ij} = 0$ , the value of  $K_{\mathcal{L}}^1$  is fixed by the constraint. For  $(i, j)$  such that  $\mathcal{L}'_{ij} \neq 0$ , any value for  $(K_{\mathcal{L}}^1)_{ij}$  other than  $(P_K)_{ij} = K_{ij}$  will result in a positive contribution in the error term, so  $(K_{\mathcal{L}}^1)_{ij} = K_{ij}$  is optimal, which concludes the proof.  $\blacksquare$

Therefore, we can decompose the matrix  $K$  into a sum of two matrices

$$K = K_{\mathcal{L}}^1 + K_{\mathcal{L}^c}^1,$$

where

$$(K_{\mathcal{L}^c}^1)_{ij} = \begin{cases} K_{ij} & : \mathcal{L}'_{ij} = 0 \\ 0 & : \text{otherwise} \end{cases}$$

*Proposition 3: EAP I with approximate cost to go*

$$\tilde{J}^1(\mathbf{z}_k) = \mathbf{z}_k^T K_{\mathcal{L}}^1 \mathbf{z}_k. \quad (7)$$

is a distributed policy, i.e.,  $\mu_i^1$  depends on  $z_i(k)$ ,  $\mathbf{z}_{-i,k}$ , and  $\mathbf{u}_{-i,k}$

*Proof:* The stage cost  $g(\mathbf{z}_k, \boldsymbol{\mu}^1(\mathbf{z}_k))$  can be expanded into

$$\begin{aligned} g(\mathbf{z}_k, \boldsymbol{\mu}^1(\mathbf{z}_k)) &= \sum_{i=1}^N g_i(z_i(k), \mathbf{z}_{-i,k}, \mu_i^1(z_i(k))) \\ g_i(z_i(k), \mathbf{z}_{-i,k}, \mu_i^1(z_i(k))) &= \frac{1}{2} \sum_{j \in \mathcal{N}_i} \kappa_C \|z_i(k) - z_j(k)\|^2 + \kappa_M \|\mu_i^1(z_i(k))\|^2. \end{aligned} \quad (8)$$

From the definition of  $K_{\mathcal{L}}^1$  and using the fact that matrix  $K$  is a positive definite  $M$  matrix, (7) can be written as

$$\begin{aligned} \tilde{J}^1(\mathbf{z}_k) &= \sum_{i=1}^N \tilde{J}_i^1(z_i(k), \mathbf{z}_{-i,k}), \\ \tilde{J}_i^1(z_i(k), \mathbf{z}_{-i,k}) &= \sum_{j \in \mathcal{N}_i} |(K_{\mathcal{L}}^1)_{ij}| \|z_i(k) - z_j(k)\|^2 + \left( (K_{\mathcal{L}}^1)_{ii} - \sum_{j \in \mathcal{N}_i} |(K_{\mathcal{L}}^1)_{ij}| \right) \|z_i(k)\|^2. \end{aligned} \quad (9)$$

In the above equation, the second summation appears because the terms corresponding to non-neighboring nodes of  $i$  are set equal to zero in the definition of  $K_{\mathcal{L}}^1$ . From Eqs. (8) and (9), it is straightforward that when solving (6), each node will effectively be solving

$$\mu_i^1(z_i(k), \mathbf{z}_{-i,k}) = \arg \min_{u_i} g_i(z_i(k), \mathbf{z}_{-i,k}, u_i) + \alpha \tilde{J}_i^1(z_i(k+1), \mathbf{z}_{-i,k+1}), \quad (10)$$

where  $z_i(k+1) = z_i(k) + Bu_i$ . To solve the above problem, node  $i$  only requires its local information, i.e., its own states and the states and inputs of its neighbors. ■

To solve (10), node  $i$  needs to know  $\mathbf{u}_{-i,k}$  as  $\tilde{J}_i^1(z_i(k+1), \mathbf{z}_{-i,k+1})$  depends on it. This can be accomplished by implementing an efficient distributed optimization algorithm that can ensure that each node finally has an accurate estimate of the control inputs,  $\mathbf{u}_{-i,k}$ , that its neighbors will implement. The algorithm that we used for simulation in the next section is a distributed subgradient algorithm that was presented in [16]. In this algorithm, each node maintains an estimate  $\hat{\mathbf{u}}_i$  of the optimization variables that it needs to compute its cost. For our problem, at time  $k$ ,  $\hat{\mathbf{u}}_i \in \mathbb{R}^{2(|\mathcal{N}_i|+1)}$  is node  $i$ 's estimate of  $u_i(k)$  and  $\mathbf{u}_{-i,k}$ . In particular, if  $\mathcal{N}_i = \{i_1, \dots, i_{|\mathcal{N}_i|}\}$ , then  $\hat{\mathbf{u}}_i = [\hat{u}_i^i \quad \hat{\mathbf{u}}_{-i}]$  where  $\hat{u}_i^j$  is node  $i$ 's estimate of  $u_j(k)$  for  $j \in \{\mathcal{N}_i \cup i\}$  and the vector  $\hat{\mathbf{u}}_{-i} = [\hat{u}_i^{i_1} \quad \hat{u}_i^{i_2} \dots \hat{u}_i^{|\mathcal{N}_i|}]$ . The main idea is the use of consensus to ensure that the estimates of the neighboring nodes converge to same values. To compute  $u_i(k)$  and  $\mathbf{u}_{-i,k}$ , node  $i$  performs the following steps:

*Distributed Projected Subgradient Algorithm [16]*

0 At iteration 0, node  $i$  initializes its estimate vector  $\hat{\mathbf{u}}_i$  with some random values.

1 At iteration  $m$ , node  $i$  updates its estimate for all  $j \in \{\mathcal{N}_i \cup i\}$  as follows:

$$v_i^j = \begin{cases} \frac{1}{2}(\hat{u}_i^j(m) + \hat{u}_j^j(m)) & : i \neq j \\ \sum_{j \in \{\mathcal{N}_i \cup i\}} \frac{1}{|\mathcal{N}_i|+1} \hat{u}_j^i(m) & : i = j \end{cases}$$

$$\hat{u}_i^j(m+1) = v_i^j - \gamma d_i(m). \quad (11)$$

2 Repeat while  $m \leq \text{iter}$ .

3  $u_i(k) = \hat{u}_i^i(\text{iter})$ .

In this algorithm, “iter” is the total number of iterations of the algorithm, and  $\gamma$  is the step size of the descent. In Eq. (11),  $d_i(m)$  is the gradient of the cost function of node  $i$  evaluated at  $\hat{\mathbf{u}}_i(m)$ . To update its estimate,  $\hat{u}_i^j(m+1)$ , node  $i$  first combines the estimates from its neighbors, which is the consensus step in Eq. (11). Then it computes the gradient of its cost at  $\hat{\mathbf{u}}_i(m)$  and updates its estimate by moving the consensus value towards the negative of the gradient.

To compute its cost to go, each node exchanges its estimates of control values with its neighbors “iter” times for all  $k$ . Although the proposed scheme has communication overhead, it is small since each node is only communicating its estimates of the current control values of itself and its neighbors. In some of the existing schemes for similar problems, nodes communicate their entire control and state trajectories with their neighbors which result in huge communication overhead depending on the horizon length and number of nodes in the network ([12] and [20]).

In EAP I, we simply imposed the sparsity structure of the communication network on the optimal cost to go matrix  $K$ . The resulting approximate terminal cost in Eq. (9) had one summation that consisted of the square of the distances of nodes  $i$  and  $j$ , such that  $(i, j) \in E$ , weighted with  $K_{ij}$ . Those were the desired terms because the objective in Problem  $\mathcal{P}$  is to minimize the distances between the neighboring nodes. However, Eq. (9) had a second summation which was

$$\left( (K_{\mathcal{L}}^1)_{ii} - \sum_{j \in \mathcal{N}_i} |(K_{\mathcal{L}}^1)_{ij}| \right) \|z_i(k)\|^2.$$

One way to interpret the above term is that each node is trying to minimize its distance from the origin. Since base station  $\mathcal{B}_1$  is assumed to be located at the origin, minimizing distance from the origin can be modeled by adding an edge between each node  $i$  and  $\mathcal{B}_1$ . Therefore, the approximate terminal cost in EAP I is with respect to a new graph  $G^I(V, E^I)$  where  $E^I = E \cup S$  where  $S = \{(i, N+1) \text{ for all } i \in \{1, 2, \dots, N\}\}$  and  $N+1$  is the index of  $\mathcal{B}_1$ . This remodeling of the structure of the system can have serious consequences that are evident in system simulation in Section IV Fig. 3. Figures 3(a) and 3(b) are the trajectories of relay nodes under the optimal policy and EAP I respectively. By comparing these trajectories, it is obvious that under EAP I, the trajectories of the nodes are biased towards the origin, which results in significant increase in the total cost of the system.

Next we propose EAP II, which introduces a refined projection of the optimal cost that removes the undesired terms from the resulting approximate cost to go. The refined projection will improve performance as will be shown

in Section IV via simulations.

### Energy Aware Policy II

Compute the positive definite matrix  $K$  offline that satisfies Riccati equation (2). Decompose the matrix  $K$  into sum of two matrices.

$$K = K_{\mathcal{L}}^{\Pi} + K_{\mathcal{L}^c}^{\Pi},$$

such that

$$(K_{\mathcal{L}}^{\Pi})_{ij} = \begin{cases} K_{ij} & : i \neq j \text{ and } \mathcal{L}'_{ij} \neq 0 \\ K_{ii} - \sum_{p \notin \mathcal{N}_i} K_{ip} & : i = j \\ 0 & : \text{otherwise} \end{cases}$$

and

$$(K_{\mathcal{L}^c}^{\Pi})_{ij} = \begin{cases} K_{ij} & : i \neq j \text{ and } \mathcal{L}'_{ij} = 0 \\ \sum_{p \notin \mathcal{N}_i} K_{ip} & : i = j \\ 0 & : \text{otherwise} \end{cases}$$

The proposed approximate cost to go function at time  $k$  is

$$\tilde{J}^{\Pi}(\mathbf{z}_k) = \mathbf{z}_k^T K_{\mathcal{L}}^{\Pi} \mathbf{z}_k,$$

where  $K_{\mathcal{L}}^{\Pi}$  is a refined projection of  $K$  on the sparsity structure of  $\mathcal{L}'$ . Each agent solves the following one stage optimization problem to compute its stationary suboptimal policy:

$$\boldsymbol{\mu}^{\Pi}(\mathbf{z}_k) = \arg \min_{\mathbf{u}} g(\mathbf{z}_k, \mathbf{u}) + \alpha \tilde{J}^{\Pi}(\mathbf{z}_k + B\mathbf{u}), \quad (12)$$

with dynamics and constraints specified in  $\mathcal{P}$ .

*Proposition 4: EAP II with approximate cost to go function at time  $k$*

$$\tilde{J}^{\Pi}(\mathbf{z}_k) = \mathbf{z}_k^T K_{\mathcal{L}}^{\Pi} \mathbf{z}_k \quad (13)$$

is a distributed policy, i.e.,  $\mu_{i,k}^{\Pi}$  depends on  $z_i(k)$ ,  $\mathbf{z}_{-i,k}$ , and  $\mathbf{u}_{-i,k}$ .

*Proof:* Using the same argument as in Prop. 3, stage cost  $g(\mathbf{z}_k, \boldsymbol{\mu}^{\Pi}(\mathbf{z}_k))$  is decentralized. For cost to go, using the definition of  $K_{\mathcal{L}}^{\Pi}$  and the fact that matrix  $K$  is a positive definite  $M$  matrix, Eq. (13) can be written as

$$\begin{aligned} \tilde{J}^{\Pi}(\mathbf{z}_k) &= \sum_{i=1}^N \tilde{J}_i^{\Pi}(z_i(k), \mathbf{z}_{-i,k}), \\ \tilde{J}_i^{\Pi}(z_i(k), \mathbf{z}_{-i,k}) &= \sum_{j \in \mathcal{N}_i} |(K_{\mathcal{L}}^{\Pi})_{ij}| \|z_i(k) - z_j(k)\|^2. \end{aligned}$$

In the above equation, the second summation consisting of the undesirable terms in Eq. (9) does not appear anymore because of the correction introduced in the definition of the refined projection matrix. It is obvious from the above arguments that when solving (12), each agent will effectively be solving

$$\mu_i^{\Pi}(z_i(k), \mathbf{z}_{-i,k}) = \arg \min_{u_i} g_i(z_i(k), \mathbf{z}_{-i,k}, u_i) + \alpha J^{\Pi}(z_i(k+1), \mathbf{z}_{-i,k+1}).$$

To solve the above problem, node  $i$  only requires its own information and the information of its neighbors. ■

Each node can compute its control action  $\mu_i^\Pi(z_i(k))$  by implementing the same distributed optimization algorithm presented for EAP I.

### B. Performance Analysis

Next we analyze the performance of the policies presented in the previous section. However, the analysis carried out in this section is for a more general system in which there are  $m$  base stations, and  $N$  mobile nodes have to establish communication links between these base stations. Furthermore, we analyze system performance for an entire class of suboptimal policies. To summarize, we are interested in the analysis of the following one step look-ahead optimization problem:

$$\begin{aligned} \min_{\mathbf{u}} \quad & g(\mathbf{z}, \mathbf{u}) + \alpha \tilde{J}(\mathbf{z}^+) \\ \text{s.t.} \quad & \mathbf{z}^+ = \mathbf{z} + B\mathbf{u}, \end{aligned} \quad (14)$$

where  $g(\mathbf{z}, \mathbf{u})$  is defined in Eq. (1). Here,

$$\tilde{J}(\mathbf{z}) = \mathbf{z}^T H \mathbf{z}$$

is an approximate cost to go,  $H$  is any symmetric positive semi-definite matrix that satisfies the constraints of communication network, and  $B = \begin{bmatrix} I_{2N} \\ 0_{2m \times 2N} \end{bmatrix}$ . Since the problem is a one stage LQR problem, the optimal policy is  $\hat{\mu}(\mathbf{z}) = -\hat{L}\mathbf{z}$ ,

$$\hat{L} = \alpha(\alpha B^T H B + R)^{-1} B^T H, \quad (15)$$

and the optimal cost for the approximate problem with one step look ahead is  $\hat{J}(\mathbf{z}) = \mathbf{z}^T \hat{K} \mathbf{z}$

$$\hat{K} = \alpha H - \alpha^2 H B (\alpha B^T H B + R)^{-1} B^T H + Q.$$

We start the analysis by proving that the optimal policy  $\hat{\mu}$  results in stabilizing system dynamics. For analysis purposes, we use the following matrix partitioning

$$U = \left[ \begin{array}{c|c} U_f & U_{fl} \\ \hline U_{lf} & U_{ll} \end{array} \right], \quad (16)$$

where  $U \in \mathbb{R}^{2(N+m) \times 2(N+m)}$ ,  $U_f \in \mathbb{R}^{2N \times 2N}$ ,  $U_{fl} \in \mathbb{R}^{2N \times 2m}$ ,  $U_{lf} \in \mathbb{R}^{2m \times 2N}$ , and  $U_{ll} \in \mathbb{R}^{2m \times 2m}$ . If  $U$  is symmetric then  $U_{lf} = U_{fl}^T$ .

**Lemma 3.1:** *Let  $H$  be a symmetric and positive semidefinite matrix, and let  $\tilde{J}(\mathbf{z}) = \mathbf{z}^T H \mathbf{z}$ . Then the system dynamics*

$$\mathbf{z}^+ = (I_{2(N+m)} - B\hat{L})\mathbf{z}$$

*are marginally stable, where  $\hat{L}$  is defined in Eq. (15).*

*Proof:* Let  $\tilde{H} = B\hat{L}$  and  $\lambda_i = \lambda_i(I_{2(N+m)} - \tilde{H})$  be the  $i^{th}$  eigenvalue of  $(I_{2(N+m)} - \tilde{H})$  such that  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{2(N+m)}$ . To prove that the system is marginally stable, we need to show that  $|\lambda_i| \leq 1$  for all

$i = \{1, 2, \dots, 2(N + m)\}$  and  $(I_{2(N+m)} - \tilde{H})$  has  $2(N + m)$  independent eigenvectors. By partitioning  $H$  as in Eq. (16),  $\alpha B^T H B = \alpha H_f$ . Thus,

$$\tilde{H} = \left[ \begin{array}{c|c} \alpha(\alpha H_f + R)^{-1} H_f & \alpha(\alpha H_f + R)^{-1} H_{fl} \\ \hline 0_{2m \times 2N} & 0_{2m \times 2m} \end{array} \right].$$

Using the properties of block matrices, the eigenvalues of  $\tilde{H}$  are the eigenvalues of  $\alpha(\alpha H_f + R)^{-1} H_f$  and  $0_{2m \times 2m}$ . Therefore,  $0_{2m \times 2m}$  contributes  $2m$  zero eigenvalues and  $\alpha(\alpha H_f + R)^{-1} H_f$  contributes  $2N$  eigenvalues. Next we will show that these  $2N$  eigenvalues are real, positive and less than one.

Since  $H$  is symmetric and positive semidefinite,  $H_f = B^T H B$  is also symmetric and positive semidefinite. Therefore,  $H_f$  has real and non-negative eigenvalues and  $2N$  independent eigenvectors. To show that the eigenvalues are also less than one, let  $\hat{\lambda}_i$  be the  $i^{th}$  eigenvalue of  $H_f$ . Then  $\frac{1}{\alpha \hat{\lambda}_i + \kappa_M}$  and  $\frac{\alpha \hat{\lambda}_i}{\alpha \hat{\lambda}_i + \kappa_M}$  are the corresponding eigenvalues of  $(\alpha H_f + R)^{-1}$  and  $(\alpha H_f + R)^{-1} \alpha H_f$  respectively. Here we have used the fact that if two matrices  $P$  and  $Q$  have the same set of eigenvectors, then they commute and  $\lambda_i(PQ) = \lambda_i(P)\lambda_i(Q)$ . Therefore, the eigenvalues of  $(\alpha H_f + R)^{-1} \alpha H_f$  are always less than or equal to one for  $\kappa_M > 0$ . This ensures that the eigenvalues of  $I_{2N} - \alpha(H_f + R)^{-1} \alpha H_f$  are also less than or equal to one. Since the last  $2m$  rows of  $\tilde{H}$  are zero,  $I_{2m} - \tilde{H}_{ll} = I_{2m}$  and there will be  $2m$  more independent eigenvectors. This implies that the eigenvalues of  $I_{2(N+m)} - \tilde{H}$  are always less than or equal to one with independent eigenvectors, which concludes the proof. ■

We can now analyze the stability properties of the proposed policies based on Lem. 3.1. Since  $K_{\mathcal{L}}^I$  is positive definite and  $K_{\mathcal{L}}^{II}$  is positive semidefinite, the dynamics for EAP I & II are stable.

An important consequence of Lem. 3.1 is that given the initial locations of the mobile relay nodes, the state space of the system is bounded since the dynamics are stable. Therefore, the performance analysis of the system can be restricted to a bounded set  $\mathcal{S}$  that is invariant under the system dynamics. We say that a set  $\mathcal{S}$  is invariant under policy  $\mu$  if  $\mathbf{z} \in \mathcal{S}$  implies that  $\mathbf{z}^+ = \mathbf{z} + B\mu(\mathbf{z}) \in \mathcal{S}$ . We define the max norm of a function  $J$  over a set  $\mathcal{S}$  by

$$\|J\|_{\mathcal{S}} = \max_{\mathbf{z} \in \mathcal{S}} |J(\mathbf{z})|.$$

For error analysis we use two mappings from [1]. Let  $\mathcal{S}_{\hat{\mu}}$  be an invariant set under policy  $\hat{\mu}$ , i.e., if  $\mathbf{z} \in \mathcal{S}_{\hat{\mu}}$  then  $\mathbf{z}^+ \in \mathcal{S}_{\hat{\mu}}$  where  $\mathbf{z}^+ = \mathbf{z} + B\hat{\mu}(\mathbf{z})$ . Let  $\mathcal{S}^*$  be the minimum set such that  $\mathcal{S}_{\hat{\mu}} \subseteq \mathcal{S}^*$  and  $\mathbf{z} \in \mathcal{S}_{\hat{\mu}}$  implies that  $\mathbf{z}^+ = \mathbf{z} + B\mu^*(\mathbf{z}) \in \mathcal{S}^*$ . For any function  $J : \mathcal{S}^* \rightarrow \mathbb{R}$ , the mappings  $T$  and  $T_{\hat{\mu}}$  are such that  $(TJ) : \mathcal{S}_{\hat{\mu}} \rightarrow \mathbb{R}$  and  $T_{\hat{\mu}} : \mathcal{S}_{\hat{\mu}} \rightarrow \mathbb{R}$  and are defined as

$$\begin{aligned} (TJ)(\mathbf{z}) &= \min_{\mathbf{u}} g(\mathbf{z}, \mathbf{u}) + \alpha J(A\mathbf{z} + B\mathbf{u}), \\ (T_{\hat{\mu}}J)(\mathbf{z}) &= g(\mathbf{z}, \hat{\mu}(\mathbf{z})) + \alpha J(A\mathbf{z} + B\hat{\mu}(\mathbf{z})). \end{aligned} \tag{17}$$

Let  $T^k$  and  $T_{\hat{\mu}}^k$  be the composition of the mappings  $T$  and  $T_{\hat{\mu}}$  with themselves  $k$  times respectively, i.e.,

$$\begin{aligned} T^k J &= T(T^{k-1} J), \\ T_{\hat{\mu}}^k J &= T_{\hat{\mu}}(T_{\hat{\mu}}^{k-1} J). \end{aligned}$$



For a stationary policy  $\hat{\mu}$ , the associated cost to go function  $J^{\hat{\mu}}$  satisfies  $\lim_{k \rightarrow \infty} T_{\hat{\mu}}^k J$ . It has been proved in [1] that the optimal cost to go  $J^*$  satisfies Bellman equation  $J^* = TJ^*$ . Similarly, for a stable policy  $\hat{\mu}$ ,  $J^{\hat{\mu}} = T_{\hat{\mu}} J^{\hat{\mu}}$ . A mapping  $T$  is a contraction mapping if there exists a scalar  $\beta < 1$  such that

$$\|TJ - T\bar{J}\|_S \leq \beta \|J - \bar{J}\|_S$$

The monotonicity lemma (Lem. 2.1 in [1]) states that for any two functions  $J$  and  $\bar{J}$  defined on  $\mathcal{S}^*$  such that

$$J(\mathbf{z}) \leq \bar{J}(\mathbf{z}) \quad \text{for all } \mathbf{z} \in \mathcal{S}^*,$$

the following inequalities hold:

$$(TJ)(\mathbf{z}) \leq (T\bar{J})(\mathbf{z}) \quad \text{for all } \mathbf{z} \in \mathcal{S}_{\hat{\mu}},$$

$$(T_{\hat{\mu}}J)(\mathbf{z}) \leq (T_{\hat{\mu}}\bar{J})(\mathbf{z}) \quad \text{for all } \mathbf{z} \in \mathcal{S}_{\hat{\mu}}.$$

Finally, for the sets  $\mathcal{S}_{\hat{\mu}}$  and  $\mathcal{S}^*$  as defined above,

$$\|T_{\hat{\mu}}\tilde{J} - TJ^*\|_{\mathcal{S}_{\hat{\mu}}} \leq \alpha \|\tilde{J} - J^*\|_{\mathcal{S}^*}.$$

To show this, the first step is to recognize that  $\hat{\mu}$  is the greedy policy with terminal cost  $\tilde{J}$ , so  $(T_{\hat{\mu}}\tilde{J}) = (T\tilde{J})$ . Let  $c = \|\tilde{J} - J^*\|_{\mathcal{S}^*}$ . Then

$$\tilde{J}(\mathbf{z}) - c \leq J^*(\mathbf{z}) \leq \tilde{J}(\mathbf{z}) + c \quad \forall \mathbf{z} \in \mathcal{S}^*$$

$$\tilde{J}(\mathbf{z}) - c \leq J^*(\mathbf{z}) \leq \tilde{J}(\mathbf{z}) + c \quad \forall \mathbf{z} \in \mathcal{S}_{\hat{\mu}}$$

The second set of inequalities hold because  $\mathcal{S}_{\hat{\mu}} \subseteq \mathcal{S}^*$ . For

$$\begin{aligned} (T(\tilde{J} - c))(\mathbf{z}) &= \min_{\mathbf{u}} [g(\mathbf{z}, \mathbf{u}) + \alpha(\tilde{J}(\mathbf{z}^+) - c)] \\ &= (T\tilde{J})(\mathbf{z}) - \alpha c. \end{aligned}$$

Similarly  $(T(\tilde{J} + c))(\mathbf{z}) = (T\tilde{J})(\mathbf{z}) + \alpha c$ . Using the monotonicity property of  $T$ ,

$$T\tilde{J}(\mathbf{z}) - \alpha c \leq (TJ)^*(\mathbf{z}) \leq T\tilde{J}(\mathbf{z}) + \alpha c \quad \forall \mathbf{z} \in \mathcal{S}_{\hat{\mu}},$$

which implies that

$$\|T_{\hat{\mu}}\tilde{J} - TJ^*\|_{\mathcal{S}_{\hat{\mu}}} \leq \alpha c = \alpha \|\tilde{J} - J^*\|_{\mathcal{S}^*} \quad (18)$$

Next we analyze the performance of any approximate policy by comparing it with the global optimal policy. We will derive bound for maximum error between the optimal and a suboptimal policy.

**Theorem 3.2:** *Let  $H$  be a symmetric and positive semidefinite matrix such that  $\tilde{J}(\mathbf{z}) = \mathbf{z}^T H \mathbf{z}$  is the approximate cost to go in (14). Let*

$$\epsilon = \|\tilde{J} - J^*\|_{\mathcal{S}^*},$$

and  $J^{\hat{\mu}} = \lim_{k \rightarrow \infty} T_{\hat{\mu}}^k \tilde{J}$ . Then the maximum error between the global optimal solution and the approximate solution is

$$\|J^{\hat{\mu}} - J^*\|_{S_{\hat{\mu}}} \leq \frac{2\alpha\epsilon}{1-\alpha} \quad (19)$$

*Proof:* The proof of this theorem is based on the properties of the mappings defined in Eq. (17).

$$\begin{aligned} \|J^{\hat{\mu}} - J^*\|_{S_{\hat{\mu}}} &= \|T_{\hat{\mu}} J^{\hat{\mu}} - J^*\|_{S_{\hat{\mu}}} \\ &\leq \|T_{\hat{\mu}} J^{\hat{\mu}} - T_{\hat{\mu}} \tilde{J}\|_{S_{\hat{\mu}}} + \|T_{\hat{\mu}} \tilde{J} - T J^*\|_{S_{\hat{\mu}}} \\ &\leq \alpha \|J^{\hat{\mu}} - \tilde{J}\|_{S_{\hat{\mu}}} + \alpha \|\tilde{J} - J^*\|_{S^*} \\ &\leq \alpha \|J^{\hat{\mu}} - J^*\|_{S_{\hat{\mu}}} + \alpha \|J^* - \tilde{J}\|_{S_{\hat{\mu}}} \\ &\quad + \alpha \|\tilde{J} - J^*\|_{S^*} \\ &\leq \alpha \|J^{\hat{\mu}} - J^*\|_{S_{\hat{\mu}}} + 2\alpha\epsilon, \end{aligned}$$

which concludes the proof. Here we have used the fact the both  $T$  and  $T_{\hat{\mu}}$  are contractions and the result proved in Eq. (18).  $\blacksquare$

One important advantage of using EAPs I & II is that they simplify the computation of these error bounds. For any general suboptimal cost  $\tilde{J}$ , it is not straightforward to compute  $\epsilon$ . However for both EAPs I & II, their corresponding values of  $\epsilon$  can easily be computed as follows:

$$\begin{aligned} \epsilon^I &= \|\tilde{J}^I - J^*\|_{S^*} = \max_{\mathbf{z} \in S^*} |\mathbf{z}^T K_{\mathcal{L}^c}^I \mathbf{z}|, \\ \epsilon^{II} &= \|\tilde{J}^{II} - J^*\|_{S^*} = \max_{\mathbf{z} \in S^*} |\mathbf{z}^T K_{\mathcal{L}^c}^{II} \mathbf{z}| \end{aligned}$$

In fact, we can derive tight upper bounds for both  $\epsilon^I$  and  $\epsilon^{II}$ . Let

$$\begin{aligned} \zeta_i &= \sum_{j \notin \mathcal{N}_i} K_{ij}, \text{ and} \\ \zeta_{\max} &= \max_{i=\{1,2,\dots,N+2\}} \zeta_i. \end{aligned}$$

Then from Gershgorin circle theorem,

$$\begin{aligned} \epsilon^I &\leq \zeta_{\max} \max_{\mathbf{z} \in S^*} \|\mathbf{z}\|^2, \\ \epsilon^{II} &\leq 2\zeta_{\max} \max_{\mathbf{z} \in S^*} \|\mathbf{z}\|^2. \end{aligned}$$

Here  $\zeta_i$  is the sum of the weights assigned to the links between node  $i$  and its non-neighboring nodes in the global optimal cost to go.

### C. Decentralized Energy Aware Policy

EAPs I & II are distributed because each node is able to compute its control action by communicating with its neighbors only. The communication overhead is small especially in the context of this problem setup in which each agent has at most two neighbors. However, in networks with dense deployment of nodes, even this

communication can cause significant energy consumption, and can result in channel congestion if all the nodes transmit simultaneously. To prevent congestion, nodes need to come up with some scheduling scheme. However, any scheduling scheme will have its own cost and will introduce latency in the system. Therefore, it is desirable to have a coordination policy that requires no inter-agent communication and can still provide some performance guarantees. We call such a policy a decentralized policy. Inter-agent communication can be avoided if the nodes are equipped with sensors that can sense the required information of the neighbors. In the absence of such sensors, a decentralized coordination policy should only require a node to communicate with its neighbors once to get their current state information. Next we propose a simple decentralized scheme that satisfies these requirements and can be implemented efficiently.

### Energy Aware Policy III

In EAP III, at time  $k$ , node  $i$  assumes that  $u_i(t) = 0$  for all  $t \geq k + 1$  and  $u_j(t) = 0$  for all  $t \geq k$  where  $j \in \mathcal{N}_i$ . Therefore, the total cost of the system is  $J(\mathbf{z}_k, \mathbf{u}_k) = \sum_{i=1}^N J_i(z_i(k), \mathbf{z}_{-i,k}, u_i(k))$ , where

$$J_i(z_i(k), \mathbf{z}_{-i,k}, u_i(k)) = \frac{1}{2} \sum_{j \in \mathcal{N}_i} \kappa_C \|z_i(k) - z_j(k)\|^2 + \kappa_M \|u_i(k)\|^2 + \frac{1}{2} \sum_{t=1}^{\infty} \alpha^t \sum_{j \in \mathcal{N}_i} \kappa_C \|z_i(k+1) - z_j(k)\|^2.$$

Since the cost is convex, the optimal control  $u_i(k)$  can easily be computed via first order necessary condition, i.e.,

$\frac{\partial J}{\partial u_i(k)} = 0$ , which yields

$$\mu_i^{\text{III}}(z_i(k), \mathbf{z}_{-i,k}) = \frac{-1}{\frac{2(1-\alpha)}{\alpha} \frac{\kappa_M}{\kappa_C} + |\mathcal{N}_i|} \sum_{j \in \mathcal{N}_i} (z_i(k) - z_j(k)).$$

This means that each agent will have consensus dynamics. The standard form of the problem under EAP III is

$$\begin{aligned} \mu^{\text{III}}(\mathbf{z}_k) &= \arg \min_{\mathbf{u}} g(\mathbf{z}_k, \mathbf{u}) + \alpha \tilde{J}^{\text{III}}(\mathbf{z}_k, \mathbf{u}), \\ \text{s.t.} \quad \mathbf{z}_{k+1} &= \mathbf{z}_k + B\mathbf{u}. \end{aligned}$$

The stage and the terminal costs are

$$\begin{aligned} g(\mathbf{z}_k, \mathbf{u}) &= \mathbf{z}_k^T Q \mathbf{z}_k + \mathbf{u}^T R \mathbf{u}, \\ \tilde{J}^{\text{III}}(\mathbf{z}_k, \mathbf{u}) &= \mathbf{z}_k^T \tilde{Q} \mathbf{z}_k + \mathbf{u}^T \tilde{R} \mathbf{u} + \mathbf{u}^T B^T \tilde{L} \mathbf{z}_k, \end{aligned} \tag{20}$$

$\tilde{Q} = c_1 \mathcal{L}'$ ,  $\tilde{R} = c_1(R' \otimes I_2)$ , and  $\tilde{L} = 2c_1 \mathcal{L}'$ . Here  $\mathcal{L}' = \mathcal{L} \otimes I_2$ ,  $R'$  is an  $N \times N$  diagonal matrix with  $R'_{ii} = |\mathcal{N}_i|$  for  $i \in \{1, 2, \dots, N\}$ , and the constant  $c_1$  is

$$c_1 = \frac{\alpha}{2(1-\alpha)} \kappa_C.$$

The resulting optimal policy in vector form is

$$\begin{aligned} \mu^{\text{III}}(\mathbf{z}_k) &= -\hat{L} \mathbf{z}_k. \\ \hat{L} &= \frac{1}{2} (R + \tilde{R})^{-1} B^T \tilde{L} \\ &= D' B^T \mathcal{L}'. \end{aligned}$$

In the above expression,  $D' = c_1(R + \tilde{R})^{-1}$ . We can also express  $D'$  as  $D' = (D \otimes I_2)$  where  $D$  is a diagonal matrix with entries

$$D_{ii} = \frac{1}{\frac{2(1-\alpha)}{\alpha} \frac{\kappa_M}{\kappa_C} + |\mathcal{N}_i|}$$

for all  $i \in \{1, 2, \dots, N\}$ .

**Lemma 3.3:** *EAP III results in a marginally stable system dynamics*

*Proof:* To show that the system dynamics are stable, we need to show that the eigenvalues of  $I_{2(N+m)} - B\hat{L}$  lie within the unit circle and its eigenvectors are independent. Using the matrix partitioning in Eq. (16), the matrix  $B^T \mathcal{L}' \in \mathbb{R}^{2N \times 2(N+m)}$  can be partitioned into  $[\mathcal{L}'_f | \mathcal{L}'_{lf}]$  and

$$I_{2(N+m)} - B\hat{L} = \left[ \begin{array}{c|c} I_{2N} - D'\mathcal{L}'_f & -D'\mathcal{L}'_{lf} \\ \hline 0_{2m \times 2N} & I_{2m \times 2m} \end{array} \right].$$

From the properties of block matrices,  $I_{2(N+m)} - B\hat{L}$  has  $2m$  eigenvalues equal to one and the remaining  $2N$  eigenvalues are the eigenvalues of  $I_{2N} - D'\mathcal{L}'_f$ . The Gershgorin circles of  $D'\mathcal{L}'_f$  are the same as those of  $D\mathcal{L}_f$  repeated twice. For  $i \in \{1, 2, \dots, N\}$ ,

$$c_i(D\mathcal{L}_f) = |\mathcal{N}_i|D_{ii} < 1 \quad r_i(D\mathcal{L}_f) \leq c_i(D\mathcal{L}).$$

This implies that  $\lambda_i(D'\mathcal{L}'_f) < 2$  and  $\lambda_i(I_{2N} - \lambda_i(D'\mathcal{L}'_f)) < 1$  for all  $i$ . Thus, all the eigenvalues of  $I_{2(N+m)} - B\hat{L}$  lie within the unit circle. To show that all the eigenvectors are independent,  $\mathcal{L}' = (\mathcal{L} \otimes I_2)$  is a laplacian matrix of an undirected graph, i.e., it is symmetric and positive semidefinite and all of its  $2N$  eigenvectors are orthogonal. The remaining  $2m$  eigenvectors are those of  $I_{2m \times 2m}$ , which are independent as well. ■

*Proposition 5:* Let the approximate cost to go  $\tilde{J}^{\text{III}}$  be as defined in Eq. (20). Let  $\epsilon^{\text{III}} = \|\tilde{J}^{\text{III}} - J^*\|_{S^*}$  and  $J^{\mu^{\text{III}}} = \lim_{k \rightarrow \infty} T_{\mu^{\text{III}}}^k \tilde{J}^{\text{III}}$ . Then the maximum error between the global optimal solution and the approximate solution is

$$\|J^{\mu^{\text{III}}} - J^*\|_{S_{\hat{\mu}}} \leq \frac{2\alpha\epsilon^{\text{III}}}{1-\alpha} \quad (21)$$

*Proof:* The proof of this proposition consists of the same steps and reasoning as the proof of Thm. 3.2. ■

#### IV. SIMULATION

In this section, we present the simulation results of the system under the proposed policies to verify their stability properties and to compare their performance with each other and with the global optimal policy. The details of the simulated system are as follows. The two base stations are separated by a distance  $d = 100$  and are located at  $[0 \ 0]^T$  and  $[0 \ 100]^T$ . The number of relay nodes is  $N = 6$  and their initial deployment locations are  $\mathbf{z}_0 = [\mathbf{x}^T \ \mathbf{y}^T]^T$ , where  $\mathbf{x} = [5 \ 12 \ 26 \ 30 \ 33 \ 38 \ 0 \ 100]^T$  and  $\mathbf{y} = [20 \ 22 \ 15 \ 17 \ 25 \ 28 \ 0 \ 0]^T$ . The last two entries in  $\mathbf{x}$  and  $\mathbf{y}$  are the locations of the base stations. The values of the length of the communication interval, communication constant, mobility constant, and the discount factor are  $T = 1000$ ,  $\kappa_C = 1$ ,  $\kappa_M = 1000$ , and  $\alpha = 0.95$  respectively. For EAPs

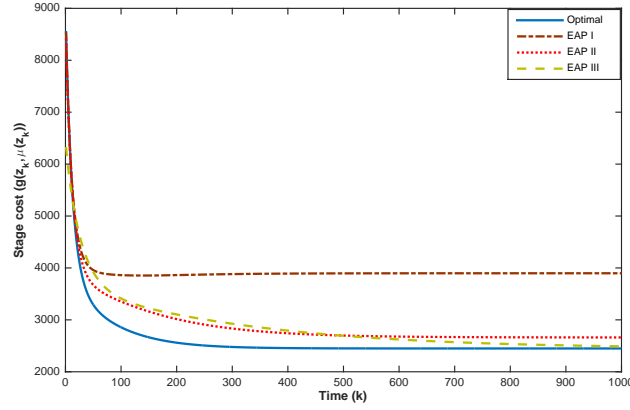
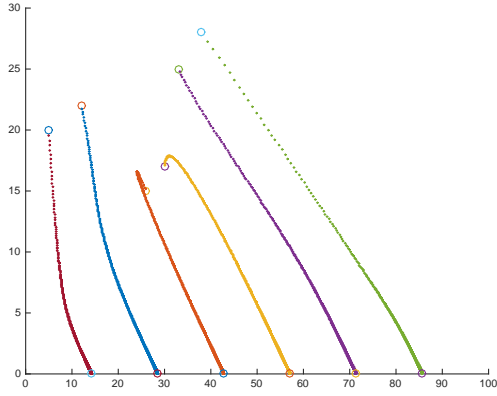


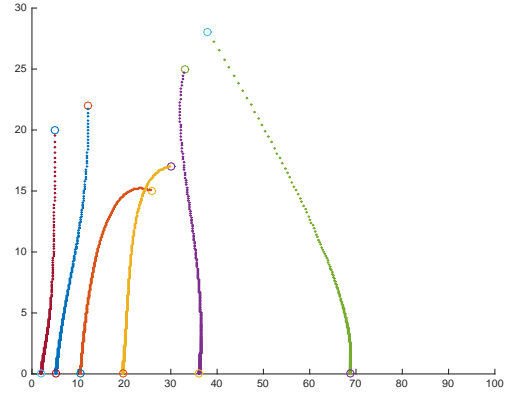
Fig. 2. Comparison of the stage costs of the proposed policies and the optimal policy.

I & II, the distributed optimization algorithm presented in Section III was implemented with number of iterations  $\text{iter} = 200$  and a fixed step size  $\gamma = 10^{-4}$ .

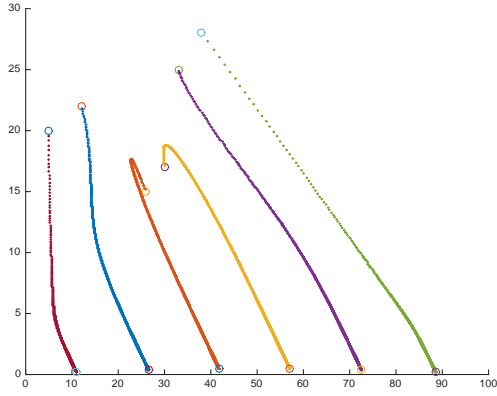
The simulation results are presented in Figs. 2 and 3. In Fig. 2, a comparison of the costs incurred by the system under the optimal and the proposed policies is presented by plotting  $g(\mathbf{z}_k, \boldsymbol{\mu}(\mathbf{z}_k))$  for  $\boldsymbol{\mu}^*$ ,  $\boldsymbol{\mu}^I$ ,  $\boldsymbol{\mu}^{II}$ , and  $\boldsymbol{\mu}^{III}$  for all  $k \in \{1, 2, \dots, T\}$ . Let the total cost incurred by the system over the interval  $[0, T]$  under policy  $\boldsymbol{\mu}$  be  $J_{\boldsymbol{\mu}}^{\text{act}} = \sum_{k=0}^T g(\mathbf{z}_k, \boldsymbol{\mu}(\mathbf{z}_k))$ . Then for the simulated system  $J_{\boldsymbol{\mu}^*}^{\text{act}} = 2.62 \times 10^6$ ,  $J_{\boldsymbol{\mu}^I}^{\text{act}} = 3.94 \times 10^6$ ,  $J_{\boldsymbol{\mu}^{II}}^{\text{act}} = 2.91 \times 10^6$ , and  $J_{\boldsymbol{\mu}^{III}}^{\text{act}} = 2.88 \times 10^6$ . Based on this comparison for this particular system, the performance of EAPs II and III is close to each other and their difference from the optimal policy is small relative to EAP I. In Fig. 3, the trajectories of the relay nodes under the optimal and the proposed policies are presented. This figure provides a good insight into the performance of the proposed policies particularly EAP I. As mentioned in Sec. III, the approximate cost to go of EAP I assumes additional edges between each relay node and the origin. Consequently, the trajectories of all the relay nodes are biased towards the origin as compared to their optimal locations. This bias towards the origin plays a fundamental role in the poor performance of EAP I. A couple of interesting observations can be made by comparing Figs. 3(c) and 3(d) with 3(a). Although the final locations of the relay nodes under EAP II are closer to the optimal terminal locations as compared to the final locations under EAP III, yet the total cost under EAP III is smaller as compared to EAP II. This observation reinforces the motivation of this work that reaching the same terminal set as the optimal solution under distributed setting may result in more energy consumption. Another interesting observation is that the relay nodes three and four first move away from their final locations and then reverse their directions. This behavior is justified because the nodes are minimizing the mobility and communication simultaneously.



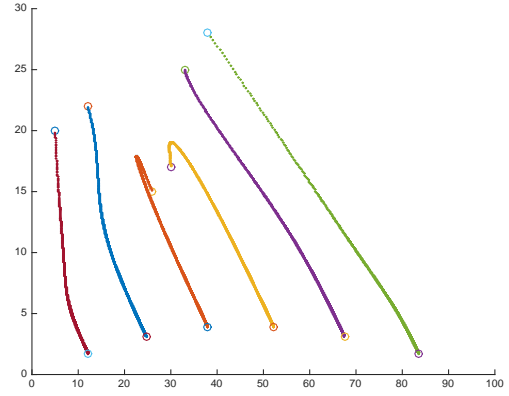
(a) Optimal trajectories



(b) Trajectories under EAP I



(c) Trajectories under EAP II



(d) Trajectories under EAP III

Fig. 3. Trajectories of the mobile relay nodes under the optimal policy and the proposed policies EAPs I, II, & III are shown in figures (a), (b), (c), and (d) respectively.

## V. CONCLUSION

We proposed an energy aware architecture for multiagent systems that can strike a balance between the global performance of the system and the cost of achieving that performance in terms of communication energy. The proposed architecture is to formulate an infinite horizon LQR problem as an equivalent one step lookahead problem with optimal cost to go as terminal cost that can be computed offline before the deployment of the system. It was shown that to compute this cost, each agent either had to communicate with all the agents in the network or communicate extensively with its immediate neighbors to compute the optimal control action, which resulted in excessive communication overhead.

*To reduce this communication overhead, the main idea behind EAPs I & II was to impose the constraints of the communication network on the global optimal cost to go and use the constrained function as an approximate cost to go. This allowed each node to compute its control action by solving a simple parameter optimization problem*

using any distributed optimization algorithm. Each node only had to exchange its current estimates of its control values and the control values of its neighbors, with its neighbors instead of exchanging estimates of all the agents as required by the globally optimal solution.

In EAP III, each node computed its control action with the assumption that all of its neighbors remain stationary. This simplifying assumption decoupled the cost of each node along the entire trajectory and allowed nodes to compute their suboptimal control action efficiently. The only information required by a node in EAP III is the current state value of its neighbors which can be either sensed if the required sensors are available or communicated.

We analyzed the performance of the proposed schemes and computed upper bounds for the performance gap between the optimal policy and the proposed policies. We also compared the performance of the proposed schemes through simulations, which showed that EAP III performed the best for the simulated system.

This setup provides a solid framework for energy aware algorithms for multiagent systems with focus on designing local interactions laws for individual nodes that are efficient in terms of energy consumption, can be implemented in real time on nodes that have limited energy and computation resources, and can provide minimum performance guarantees.

#### APPENDIX

*Proposition 1: If  $Q \in \mathbb{R}^{2(N+2) \times 2(N+2)}$  is a symmetric  $M$  matrix,  $B = \begin{bmatrix} I_{2N} \\ 0_{4 \times 2N} \end{bmatrix}$  and  $R = \kappa_M I_{2N}$ , then the positive definite solution of the Riccati equation*

$$K = \alpha K - \alpha^2 K B (\alpha B^T K B + R)^{-1} B^T K + Q.$$

*is also an  $M$  matrix. Moreover, if  $Q$  is a laplacian matrix of a connected graph then all the entries of  $K$  are non-zero.*

*Proof:* If  $B = I_{2(N+2)}$  and  $\alpha = 1$ , then this result has been proved in Theorem 5.1 of [19]. We will extend this result for  $B$  and  $\alpha$  as defined above. We start the proof by representing the matrices  $Q$  and  $K$  as block matrices following the convention introduced in (16):

$$Q = \left[ \begin{array}{c|c} Q_f & Q_{fl} \\ \hline Q_{fl}^T & Q_l \end{array} \right] \text{ and } K = \left[ \begin{array}{c|c} K_f & K_{fl} \\ \hline K_{fl}^T & K_l \end{array} \right],$$

Then,

$$\begin{aligned} Q &= (1 - \alpha)K + \alpha^2 K B (\alpha B^T K B + R)^{-1} B^T K. \\ B^T Q B &= (1 - \alpha)B^T K B + \alpha^2 B^T K B (I + \alpha B^T K B R^{-1})^{-1} R^{-1} B^T K B \\ R^{-1} Q_f &= (1 - \alpha)R^{-1} K_f + \alpha^2 R^{-1} K_f (I + \alpha K_f R^{-1})^{-1} R^{-1} K_f \end{aligned} \tag{22}$$

The last equality is based on matrix decomposition,  $B^T K B = K_f$  and  $B^T Q B = Q_f$  and multiplying both sides

by  $R^{-1}$ . Because  $R = \kappa_M I_{2N}$ , it commutes with any square matrix. Using matrix inversion lemma

$$\begin{aligned} R^{-1}Q_f &= R^{-1}K_f - \alpha(I + \alpha K_f R^{-1})^{-1}R^{-1}K_f \\ (R^{-1}K_f)(R^{-1}Q_f) &= (R^{-1}K_f)^2 - \alpha R^{-1}K_f(I + \alpha K_f R^{-1})^{-1}R^{-1}K_f \end{aligned}$$

From Eq. (22),  $\alpha^2 R^{-1}K_f(I + \alpha K_f R^{-1})^{-1}R^{-1}K_f = R^{-1}Q_f - (1 - \alpha)R^{-1}K_f$ . Using this equality in the above equation yields the following quadratic equation.

$$(R^{-1}K_f)^2 - R^{-1}K_f \left( R^{-1}Q_f - \frac{1 - \alpha}{\alpha} I \right) - \frac{1}{\alpha} R^{-1}Q_f = 0$$

After performing a series of simple algebraic manipulations

$$R^{-1}K_f = G + \sqrt{G + \frac{4}{\alpha}\sqrt{G}} - \frac{2(1 - \sqrt{\alpha})}{\alpha} I_{2N},$$

where  $G = R^{-1}Q_f + \frac{(1 - \sqrt{\alpha})^2}{\alpha}$ . Since  $Q_f$  is a symmetric positive semidefinite  $M$  matrix,  $G$  is a positive definite  $M$  matrix and  $G_s = \sqrt{G}$  is also a positive definite  $M$  matrix [21], and all of its entries are non-zero [19]. Using Lemma 5.5 of [19],  $G + \sqrt{G + \frac{4}{\alpha}\sqrt{G}}$  is an  $M$  matrix with negative off-diagonal entries. Now,  $R^{-1}K$  is a positive definite matrix, so its principal submatrix  $R^{-1}K_f$  is also positive definite. Thus, all the diagonal entries of  $R^{-1}K_f$  are positive proving that  $K_f$  is a positive definite  $M$  matrix.

Next we will show that all the entries of  $K_{fl}$  are negative. Using the definition of  $B$ ,

$$\begin{aligned} Q_{fl} &= (1 - \alpha)K_{fl} + \alpha^2 K_f (\alpha K_f + R)^{-1} K_{fl} \\ K_{fl} &= [(1 - \alpha)I + \alpha^2 K_f (\alpha K_f + R)^{-1}]^{-1} Q_{fl} \\ &= (1 - \alpha)(sI - D)^{-1} Q_{fl} \end{aligned}$$

where  $s = \left(1 + \frac{\alpha^2}{1 - \alpha}\right)$  and  $D = \frac{\alpha^2}{1 - \alpha}(I + \alpha R^{-1}K_f)^{-1}$ . Because  $\alpha R^{-1}K_f$  is an  $M$  matrix,  $D$  is a positive matrix. Furthermore,  $\rho((I + \alpha R^{-1}K_f)^{-1}) < 1$ , which implies that  $\rho(D) < s$  and  $(1 - \alpha)(sI - D)^{-1}$  is a positive matrix. Since  $Q_{fl}$  has all the entries non-positive,  $K_{fl}$  has all the entries strictly negative unless an entire column of  $Q_{fl}$  is zero which is not possible for a connected graph.

Finally, to show that  $K_l$ ,

$$Q_l = (1 - \alpha)K_l + \alpha^2 K_{fl}^T (\alpha K_f + R)^{-1} K_{fl}$$

Since  $(\alpha K_f + R)^{-1}$  is a positive matrix and all the entries of  $K_{fl}$  are negative,  $\alpha^2 K_{fl}^T (\alpha K_f + R)^{-1} K_{fl}$  is a positive matrix. By definition  $Q_l$  has positive diagonal entries and non-positive off-diagonal entries. To satisfy the above equation,  $K_l$  must have positive diagonal entries and negative off-diagonal entries. This concludes the proof. ■

## REFERENCES

- [1] D. P. Bertsekas and J. N. Tsitsiklis, *Neuro-Dynamic Programming*, Athena Scientific, 1996.
- [2] J. R. Marden and J. S. Shamma, "Game Theory and Distributed Control," in *Handbook of Game Theory*, ( H. P. Young and S. Zamir, Eds.), vol. 4, Elsevier Science, 2014.



- [3] G. Anastasi, M. Conti, M. Francesco, and A. Passarella, "Energy conservation in wireless sensor networks: A survey," in *Ad Hoc Networks*, vol. 7, no. 3, pp. 537–568, 2009.
- [4] M. Zhu and S. Martinez, "Distributed coverage games for energy-aware mobile sensor networks," in *SIAM J. Control Optim.*, vol. 51, iss. 1, pp. 1–27, 2012.
- [5] D.K. Goldenberg, J. Lin, and A.S. Morse, "Towards Mobility as a Network Control Primitive," in *Proc. ACM MobiHoc*, pp. 163–174, 2004.
- [6] F. El-Moukaddem, E. Torng, G. Xing, and S. Kulkarni, "Mobile relay configuration in data-intensive wireless sensor networks," in *IEEE Transactions on Mobile Computing*, vol. 12, no. 2, pp. 261 - 273, Feb. 2013.
- [7] H. Jaleel, Y. Wardi, and M. Egerstedt, "Minimizing mobility and communication energy in robotic networks: An optimal control approach," in *Proc. ACC*, Portland, OR, June 2014.
- [8] U. Ali and Y. Wardi, "Multiple Shooting Technique for Optimal Control Problems with Application to Power Aware Networks," in *Analysis and Design of Hybrid Systems*, Atlanta, GA, USA, Oct. 2015.
- [9] G. Ferrari-Trecate, L. Galbusera, M. P. E. Marciandi, and R. Scattolini, "Model predictive control schemes for consensus in multi-agent systems with single and double integrator dynamics," in *IEEE Trans. Autom. Control*, vol. 54, iss. 11, pp. 2560–2572, Nov. 2009.
- [10] M. A. Müller, M. Reble, and F. Allgöwer, "Cooperative control of dynamically decoupled systems via distributed model predictive control," in *Int. J. of Robust. & Nonlinear Control*, vol. 22, pp. 1376–1397, 2012.
- [11] H. Li, W. Yan, "Receding horizon control based consensus scheme in general linear multi-agent systems," in *Automatica*, vol. 56, pp. 12–18, 2015.
- [12] W. B. Dunbar and R. M. Murray, "Distributed receding horizon control for multi-vehicle formation stabilization," in *Automatica*, vol. 42, no. 4, pp. 549–558, Apr. 2006.
- [13] E. Semsar-Kazerooni, K. Khorasni, "Optimal consensus algorithms for cooperative team of agents subject to partial information," in *Automatica*, vol. 44, pp. 2766–2777, 2008.
- [14] D. Bauso, L. Giarre, R. Pesenti, "Non-linear protocols for optimal distributed consensus in networks of dynamic agents," in *Systems & Control Letters*, vol. 55, pp. 918–928, 2006.
- [15] R.L. Raffard, C.J. Tomlin, and S.P. Boyd, "Distributed optimization for cooperative agents: Application to formation flight," in *Proc. 43rd IEEE Conf. Dec. Control*, pp. 2453–2459, 2004.
- [16] A. Nedic, A. Ozdaglar, and P. Parrilo, "Constrained consensus and optimization in multi-agent networks," in *IEEE Trans. Autom. Control*, vol. 55, no. 4, pp. 922–938, 2009.
- [17] C. R. Johnson, "Inverse M-matrices," in *Linear Algebra Appl.* vol. 47, pp. 195–216, 1982.
- [18] Y. Mei, Y-H. Lu, Y. C. Hu, and C.S. G. Lee, "A case study of mobile robot's energy consumption and conservation techniques," in *Proc. 12<sup>th</sup> IEEE ICAR*, pp. 492–497, 2005.
- [19] Y. Cao and W. Ren, "Optimal linear consensus algorithms: An LQR perspective," in *IEEE Trans. System, Man, Cybernetics*, vol. 40, no. 3, pp. 819–830, Mar. 2010.
- [20] T. Keviczky, F. Borrelli, and G. J. Balas, "Decentralized receding horizon control for large scale dynamically decoupled systems," in *Automatica*, vol. 42, no. 12, pp. 2105–2115, Dec. 2006.
- [21] G. Alefeld and N. Schneider, "On square roots of M-matrices," in *Linear Algebra Appl.*, vol. 42, pp. 119–132, Feb. 1982.